人工智能与人类智能
——从认知科学五个层级的理论看人机大战

蔡曙山[1]　薛小迪[2]

(1. 贵州民族大学 民族文化与认知科学学院,贵州 贵阳 550000;2. 清华大学 心理学系,北京 100084)

摘　要:人类心智和认知分为神经、心理、语言、思维、文化五个层级。在人类认知的所有五个层级上,人工智能都是在模仿人类智能;人工智能是在不断进步的,但在总体上并未超过人类智能。在语言、思维和文化层级上,即在高阶认知层级上,目前人工智能都远逊于人类智能。事实上,人工智能和人类智能这两种智能方式是截然不同的。机器学习也只是对人类认知能力的一种模仿,不能作过高的评价,不必惊慌,更不能人为地制造恐慌。著名的塞尔人工智能理论"中文房间论证"的模型和标准并未过时,根据这个论证,数字计算机(digital computer)永远不会具有人类智能,因此,在可预见的未来也不会出现超过人类智能的人工智能。

关键词:人类心智;人类认知;人类智能;人工智能;认知科学

中图分类号:B 842.1　　文献标识码:A　　文章编号:1000-5919(2016)04-0145-10

最近牵动学界、引起国家决策层关注的人机大战(李世石 PK 谷歌狗 AlphaGo)①,各种看法、意见和评论莫衷一是,其中不乏真知灼见,但很多看法却不得要领,更有一些看法有意无意夸大机器智能的作用和意义,甚至宣扬人类历史的终结和世界末日论,这当然是不足为训。我们认为最根本的就是不能正确理解人工智能与人类智能,特别不能从人类所特有的语言、思维、文化层级上来正确地理解人工智能与人类智能。

按照蔡曙山(2015,2016)对人类心智和认知五个层级划分的理论②,我们应该而且能够从神经、心理、语言、思维、文化五个层级来认识和区分机器智能(人工智能)与人类智能,为当前正在发生的人工智能 PK 人类智能找到正确的解释和答案。

一、人类认知的五个层级

图1是众所周知的认知科学的学科结构图。③

收稿日期:2016-04-15

作者简介:蔡曙山,男,贵州贵阳人,清华大学心理学系、贵州民族大学民族文化与认知科学学院教授。

基金项目:国家社会科学基金重大项目"语言、思维、文化层级的高阶认知研究"(项目批准号:15ZDB017),"汉语非字面语言大脑加工的神经机制研究"(项目批准号:14ZDB154)。

① 2016年3月9—15日,李世石与 AlphaGo 的世纪大战,李世石以1比4总成绩负于 AlphaGo。事件不仅给学界带来极大震荡,也惊动了国家相关部门和中央决策层。3月15日当天,中央办公厅和教育部要求人工智能相关领域专家就此事件提供咨询意见,回答人工智能是什么、怎么看、怎么办的问题,供领导决策参考。本文是在笔者提交的咨询报告上增写而成的,以认知科学的原理和笔者建立的人类认知五层级理论为依据,回答了人工智能与人类智能本质区别等一列重大理论问题。

② 蔡曙山:《论人类认知的五个层级》,《学术界》2015年第12期。《人类认知的五个层级和高阶认知》,《科学中国人》2016年第4期。

③ Pylyshyn, Z. Information science: its roots and relations as viewed from the perspective of cognitive science. In Machlup, F., and Mansfield, U. (eds.) (1983), *The Study of Information: Interdisciplinary Messages*, New York: Wiley.

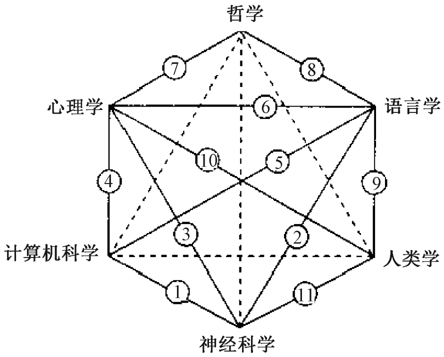

图 **1**　认知科学的学科结构(Pylyshyn,**1983:76**)

这个学科结构图展示的是认知科学的学科结构和学科关系,但它却并不展示认知科学的研究对象及对象间的关系,因为认知科学的对象并不是这些学科,不是哲学,不是心理学,不是语言学,不是计算机科学,不是人类学,也不是神经科学——所有这些学科都不是认知科学的对象。认知科学的对象是人类的心智和认知。那么,人类的心智和认知又具有一种什么样的内涵和关系呢?

2015 年以来,笔者著文提出人类心智和认知五个层级的理论。根据人类心智进化的历程,人类心智从初级到高级可以分为五个层级:神经层级的心智、心理层级的心智、语言层级的心智、思维层级的心智和文化层级的心智。由于认知是用心智来定义的①,故人类认知从初级到高级也同样分为五个层级:神经层级的认知、心理层级的认知、语言层级的认知、思维层级的认知和文化层级的认知,简称神经认知、心理认知、语言认知、思维认知和文化认知。五个层级的认知是人类心智进化各个阶段认知能力的存留。人类认知只能而且必需被包含在这五个层级之中。神经认知和心理认知是人和动物共有的,称为“低阶认知”(lower-order cognition),语言认知、思维认知和文化认知是人类所特有的,称为“高阶认知”(higher-order cognition)。五个层级的认知形成一个序列,他们的关系如图 2。在这个序列中,低层级的认知是高层级认知的基础,或者说,低层级的认知决定高层级的认知,而高层级的认知向下包含并影响低层级的认知。②

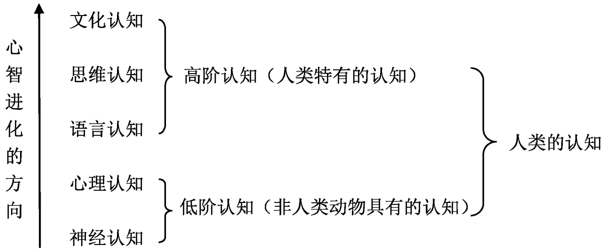

图 **2**　人类认知的五个层级

在人类认知五个层级中,语言认知处于非常特殊的地位。其一,语言区分了人类认知和动物认知。从进化的眼光看,直立行走、火的使用、语言的发明这三件事改变了进化的方向,使人最终进化成人。直立行走使人的活动范围大大扩展,从丛林走向平原,解放的前肢用于采摘和捕猎,扩大了食物的范围。火的使用使人可以吃其他动物的肉,异体蛋白的大量摄入使人的脑容量空前增大,脑的认知能力大大提高。语言的发明是人类进化关键的一步。自从使用表意的符号语言和文字,人类的经验就可以形成知识,积淀为文化,从此,人类的进化不再是动物的基因层级的进化,而是语言、知识和文化层级的进化。

① 蔡曙山:《认知科学框架下心理学、逻辑学的交叉融合与发展》,《中国社会科学》2009 年第 2 期。

② 蔡曙山:《论人类认知的五个层级》,《学术界》2015 年第 12 期。

其二,语言使思维成为可能。人类的语言能力表现在,它主要通过隐喻的方法,产生和使用抽象概念,并在抽象概念的基础上,形成判断,进行推理。应用判断和推理,人类可以进行决策和丰富多彩的思维,包括数学思维、物理学思维、哲学思维、文学思维、历史思维、艺术思维等等。人类社会的一切,都是应用语言和思维的结果。法国著名数学家、逻辑学家和哲学家笛卡尔有一句名言“我思,故我在”——人类的思维是人类存在的原因。20 世纪最重要的语言和思维关系的理论假说“沃尔夫假说”(B. L. Whorf,1956)是语言形成思维的观点,它由两个部分构成:一是语言决定论(linguistic determinism),指语言决定非语言过程,即学习一种语言会改变一个人思维的方式;二是语言相对性(linguistic relativity),指被决定的认知过程因语言不同而不同,因此,不同语言的说话人以不同的方式进行思维。沃尔夫假说在人类五个层级认知的理论框架下得到完美的和完全的解释。

其三,语言和思维形成知识,知识积淀为文化。非人类的动物则只能由每一代和每一个体重新开始积累经验,其进化只能是基因层次的进化。人类知识绝大部分来源于前人创造和积累的间接知识,其进化不仅仅是基因层次的进化,更主要的是知识的进化。自从人类发明和使用语言文字,人类的历史可以说是日新月异,而在此之前,人类的进化与其他动物一样,是以千、万年为单位的。①

其四,“语言的限度就是我的世界的限度。”20 世纪西方哲学发展前后相继的三大主流是分析哲学、语言哲学和心智哲学。其代表性人物有维特根斯坦、乔姆斯基、奥斯汀和塞尔等。维特根斯坦前后两个时期的两本重要著作《逻辑哲学论》(1921)和《哲学研究》(1953)奠定了分析哲学和语言哲学的基础。事实上,这两部划时代的著作以不同方式研究了语言的哲学问题。维特根斯坦说:“语言的述说乃是一种活动,或是一种生活形式的一个部分。”②“一个词的意义就是它在语言中的使用。”③“哲学是一场反对借我们的语言来迷惑我们的理智的战斗。”④他还说:“全部哲学都是一种‘语言批判’。”⑤“我的语言限度就是我的世界限度。”⑥因此,“对于不可言说的东西,我们必须保持沉默”。⑦

其五,语言建构社会现实。乔姆斯基的《句法结构》(1957)、蒙太格的《形式哲学》(1974)、奥斯汀的《如何以言行事》(1962)分别开创了当代语言学研究的语形学(syntax)、语义学(semantics)和语用学(pragmatics)的研究。这个三分框架,是现代语言研究和语言学的基本方法,也为心理学、社会学、政治学、艺术学等众多学科所使用。⑧ 在语言哲学和心智哲学领域贡献卓著的人物是世界著名语言和心智哲学家塞尔(John R. Searle)。塞尔的《言语行为》(1969)、《表达式和意义》(1979)、《意向性》(1983)、《语用逻辑基础》(1985)、《意识之谜》(1997)、《建构社会现实》(1997)、《心智、语言和社会》(1999)、《心智》(2005)、《制造社会世界》(2010)等,建立言语行为理论和语用逻辑、意向性理论、意识理论、心智和认知理论、语言建构社会现实理论等一系列关于语言、心智和认知的重要理论。在语言建构社会现实的理论方面,塞尔在言语行为理论、语用学、语言哲学和心智哲学基础上,提出人类用语言建构整个人类社会,这一理论振聋发聩,引起学界和社会各界广泛关注。笔者曾与应邀访问清华大学并主持清华论坛的塞尔教授有过关于语言建构社会现实的对话。塞尔说:

---

① 人类言语和语言的使用和文字的发明对人类认知能力的影响,参阅米黑尔·罗科、威廉·班布里奇编,蔡曙山等译:《聚合四大科技,提高人类能力:纳米技术、生物技术、信息技术和认知科学》,北京:清华大学出版社 2010 年版,第 32 页表 3。

② 维特根斯坦:《哲学研究》,李步楼译,北京:商务印书馆 2004 年版,第 17 页。

③ 维特根斯坦:《哲学研究》,李步楼译,第 31 页。

④ 维特根斯坦:《哲学研究》,李步楼译,第 71 页。

⑤ 维特根斯坦:《逻辑哲学论》,贺绍甲译,北京:商务印书馆 2009 年版,第 42 页。

⑥ 维特根斯坦:《逻辑哲学论》,贺绍甲译,第 85 页。

⑦ 维特根斯坦:《逻辑哲学论》,贺绍甲译,第 105 页。

⑧ 蔡曙山:《符号学三分法及其对语言哲学和语言逻辑的影响》,《北京大学学报》2006 年第 5 期。

> 从社会本体论的意义上说,人类社会是由语言建构的,并且不断传递下去。正如生物学的 DNA 一样,人类社会也有普遍的原则,而这些原则正是用语言来建构的。……我们确实不知道区别人类心智与动物心智的细节,但我们知道人类语言具有与动物语言不同的特征。动物的语言可以用于表达,人类的语言却可以用于表现,即用于建构社会现实。①

——从五个层级的理论和以上分析我们看出,语言是人类心智的基础,语言决定思维和我们认识世界的方式,语言的限度就是我的世界的限度。人类无法认识世界,也无法做任何事情——除非经过语言。

二、认知科学家眼中的计算机科学与人工智能

关于认知科学,现在我们有两个关系图。其一是“学科关系图”,即图 1,它是从认知科学的来源学科来划分的,依据的是学科标准,说明的是认知科学各来源学科和交叉学科之间的关系。其二是“科学关系图”,即图 2,这是从人类认知的五个层级来划分的,依据的是人们头脑里发生的认知过程,说明的是认知科学研究对象之间的关系。

我们注意到,在认知科学的学科关系图中,作为认知科学 6 大来源学科之一的计算机科学是赫然在目的,并且作为计算机科学与认知科学交叉的 6 大支柱学科的人工智能也是认知科学的主流学科。但是,在认知科学的五个层级的划分中,即在认知科学的科学结构中,计算机科学和人工智能都没有了踪影,这是怎么回事?

首先,我们应该看到,在科学和学科的关系中,科学是第一性的,是决定的;学科是第二性的,是被决定的。科学和学科的这种关系,在认知科学五个层级的科学结构和六大学科的学科结构关系中,得到充分的体现(见表 1)。

表 **1** 认知科学的科学结构和学科结构对应关系

层级	认知形式	问题和领域	学科
5	文化认知	自我、他人、社会、文化、自然、进化	文化学;人类学;文化人类学
4	思维认知	概念、判断、推理、证明、决策、问题解决	逻辑学;哲学;计算机科学
3	语言认知	句法加工、语义加工、语用加工	语言学
2	心理认知	感觉、知觉、注意、表象、记忆	心理学
1	神经认知	视觉、听觉、触觉、嗅觉、味觉	神经科学

从以上分析我们可以看出:第一,科学关系决定学科关系。与其他科学研究一样,认知科学也是以问题为导向的,认知科学的五个层级的研究所涉及的问题分别产生出神经科学、心理学、语言学、逻辑学和哲学、计算机科学、文化人类学。科学与学科的对应关系如下:神经认知→神经科学;心理认知→心理学;语言认知→语言学;思维认知→逻辑学,哲学,计算机科学;文化认知→人类学,文化人类学。20 世纪 50 年代以来,五个层级的认知研究最终汇成大海,诞生了认知科学。所以,科学问题和科学研究是第一性的,是决定的;学科的产生和发展是第二性的,是被决定的。认知科学也是一样,其科学性质决定学科性质,即是说,认知科学的五个层级的科学关系决定其六大学科的结构和相互关系。事实上,图 1 的认知科学学科结构是从图 2 的认知科学五个层级通过科学到学科的映射而得的。遗憾的是,长期以来,在认知科学研究和学科建设中,人们并未充分重视比学科结构更加基本的科学结构。第二,思维认知到认知科学相关学科的映射是一对多的关系,产生了逻辑学、哲学、计算机科学等相关学科。这充分说明,思维是人类认知的重要形式,所以对它的研究形成多个学科。从思维层级的认知到认知科学相关学科

① 马欣:《塞尔:人类语言用于建构社会现实》,《科学时报》2007 年 9 月 11 日。

的映射,充分说明了思维认知的重要性。第三,计算机并不是人脑或人类心智的一部分,但它却是人类心智和认知的产物,并反过来促进人类心智和认知的发展,这就形成了人工智能与人类智能的对立和统一。人工智能和人类智能的统一性表现在两个方面:其一,计算机科学和人工智能是人类心智和认知的一种外在的形式和工具。其二,计算机科学和人工智能作为一门学科,它的对象存在于人类认知的五个层级之中。

从上述的认知科学五个层级到认知科学的学科之间的映射关系我们看出,计算机和人工智能不过就是试图用机器来实现人类智能。

要实现人工智能的这个目标,首先就要了解什么是人类智能,然后才能正确理解什么是人工智能。

人类智能,就是神经、心理、语言、思维、文化五个层级上所体现的人类的认知能力。

人工智能,就是让机器或人所创造的其他人工方法或系统来模拟人类智能。

关于人工智能,有两个重要的标准或模型,一个是图灵模型,由英国数学家和逻辑学家、计算机和人工智能之父图灵(Alan Turing)于 1948 年在“智能机器”中提出;另一个是塞尔模型,由美国心智与语言哲学家塞尔(John R. Searle)于 1980 年提出。

图灵模型是一部由人操控的机器,它按照用自然语言(英语)书写的程序来下棋,是一部“纸上谈兵”的机器(a paper machine)。操控这部机器的人无须知道如何下棋,他所做的是只需按照书写的程序去移动棋盘上的棋子。在人工智能上图灵是一位乐观主义者,他相信计算机很快就会展现明显的智能行为,例如回答英文提出的问题并进行对话等等。1950 年,图灵提出著名的“图灵试验”(Turing Test):如果一台计算机能够通过人的在线会话测试,则这台机器就被认知已经具有智能。与此等价的模型是:如果一位测试者在一场屏蔽对象的在线会话测试中不能分辨由人和机器分别控制的两个终端的回答哪一个更好,则说明被测试的机器已经具有了智能。按照这个标准,很显然,目前的很多计算机系统已经具有了智能,如 1997 年击败卡斯帕罗夫的“深蓝”、最近击败李世石的 AlphaGo,以及在各种专业领域中其智能堪比人类的各种专家系统。到 20 世纪 70 年代晚期,一些人工智能研究者宣称计算机已经理解至少部分自然语言。1980 年,美国加州大学伯克利分校的心智和语言哲学家,美国科学与艺术学院院士、美国人文社科总统奖章获得者塞尔,以一个简明而被广泛讨论的模型,断言当前的数字计算机完全不可能理解自然语言或人类思维。1999 年,塞尔简明地描述其“中文房间论证”(Chinese Room Argument)模型如下:

> 设想一个母语为英语的人,他对汉语一无所知,他被锁在一个装有中文字符盒子的房间中,房间中还有一本关于中文字符操作的指导手册(程序)。又设想房间外面的人往房间里送进一些中文字符,房间里的人对这些字符依然是不认识,而这些送进来的字符是以中文提出的问题(输入)。再设想房间里的人按照程序指导能够发出中文字符,而这些中文字符也正确回答了那些问题(输出)。程序使房间中的人通过了图灵试验,然而,房间里的这个人确实对中文一无所知。①

塞尔继续说道:“这一论证的要点是:如果中文房间中的人通过操作适当的程序来理解中文,但他却并不理解中文,那么,任何仅仅基于同样程序的数字计算机也是不理解中文的,因为中文房间里的人所不具有的东西,任何计算机作为计算机也不可能具有。”②

——这就是认知科学家所看待的人工智能。

下面,我们从人类智能的五个层级来分析人工智能,看看目前的人工智能包括 AlphaGo 都达到了什

---

① The Chinese Room Argument, *Stanford Encyclopedia of Philosophy*, http://plato. stanford. edu/entries/ chinese - room/。

② The Chinese Room Argument, *Stanford Encyclopedia of Philosophy*, http://plato. stanford. edu/entries/ chinese-room/。

么水平,和人类智能究竟有什么差异?

三、以五个层级的理论来认知人工智能和人类智能

(一)神经层级的认知

在神经认知层级上,计算机科学与神经科学交叉产生了两个重要的科学领域:神经计算机科学和计算神经科学。计算机科学和人工智能目前在神经层级所做的工作不外这两个领域。那么,在这些领域,人工智能是否已经超过或至少相当于人类智能了呢?答案是否定的。

AlphaGo 目前所做的工作,只是模仿人类神经的某些活动,却取了一些相当吓人的名称,如"神经网络""神经计算机""类脑计算机""深度心智"等等,其实与人类的神经认知活动一点关系也没有。人类的神经认知活动,如视觉认知、听觉认知、嗅觉认知、味觉认知、触觉认知,计算机和人工智能远没有达到人类的认知能力和水平;而对幸福、痛苦和各种情绪的感受,目前的人工智能恐怕连一些低级的动物如虫鱼鸟兽的认知水平都比不上。

(二)心理层级的认知

在心理认知层级上,计算机科学与心理学交叉的科学领域有计算机仿真和计算心理学等。在这些领域,人工智能是否已经超过或相当于人类智能了呢?答案仍然是否定的。

以感知觉和注意、表象和记忆这些基本的心理现象为例,计算机和人工智能目前远没有达到人类的认知水平。例如,视觉中对颜色的感觉即色觉可以兼有温度感觉,如红、橙、黄色会使人感到温暖,所以这些颜色被称作暖色;蓝、青、绿色会使人感到寒冷,这些颜色被称作冷色等等。计算机的视知觉系统未能实现这种跨通道的感知。又如,知觉是脑和神经系统对感觉信息再加工,以获得对事物的整体性认识的心理过程。知觉具有整体性、恒常性、意义性、选择性等特征。一辆自行车,我将它靠在一棵树上,回来时从另一个方向和完全不同的角度看它,一眼就能认出那是我的自行车,而机器认知系统或人工智能系统对此却可能一筹莫展。

(三)语言层级的认知

计算机的语言系统与人类的语言系统有本质的不同。人类语言是在自然进化中形成的语言,称为自然语言。计算机语言却是人类自己设计出来的、专供机器使用的语言,它是人工语言中的形式语言。

自然语言和形式语言有巨大的差异。以汉语为例,以汉语为母语的中国人很容易知道下面的话语哪些是可以说的,哪些是不可以说的:

(a)吃饭。

(b)吃酒席。

(c)吃食堂。

(d) * 吃桌子。

(e) ** 吃教室。

(f) ** 吃操场。

注意在现代汉语中,(a)(b)(c)是可以说的,(e)和(f)却是不可以说的,(d)在一些语境中是可以说的,在另一些语境中却是不可以说的。母语是汉语的中国人很容易掌握这种语言知识和技能,而计算机理解起来却非常困难。原因是人类在使用自然语言时,可以同时在句法、语义和语用三个层次进行加工,但目前的计算机自然语言处理系统在最基础的句法加工上的表现也只能是差强人意(仅对英语而言),在语义加工和语用加工上则只能是"望文兴叹"或"听语兴叹"。

计算机科学和语言学交叉的领域是计算语言学和自然语言理解。这些领域可能的突破性进展还要依赖于认知科学的参与。人类认知或者说高阶认知的基础是语言认知,思维和文化认知都是建立在语言认知基础之上的(详见本文第一部分)。人类凭借生动活泼、丰富多彩的自然语言来进行思维,并形

成瑰丽多彩的人类文化。计算机则凭借单调的、无歧义的二进制语言来进行一切活动,包括对人类思维的模拟。形式语言与自然语言之间的鸿沟是一道难以跨越的障碍,而由不同的语言所形成的这种障碍是人工智能和人类智能之间的根本分野。

(四)思维层级的认知

人工智能与人类智能的重要区别,也许就在思维层级的认知上。

计算机是否有智能,这个问题常常被当作计算机是否能够思维这个问题——这样的理解存在重大的偏差,尽管思维是人类认知的重要方面。

人类的心智和认知是以语言为基础,以思维为特征的。思维是人类做出的最高级别的精神活动。所有人类的进步和成就不过就是人类思想的产物。文化、艺术、科学和技术的发展无一不是思维的结果。① 甚至整个人类社会都是用语言和思维来建构的。法国哲学家笛卡尔的著名论断"我思故我在"将思维与存在的关系定义为因果关系:由于我思维,所以我存在。笛卡尔的心身问题(Mind-Body problem)是哲学和认知科学的永恒问题。②

中国古代思想家对于思维也有过非常精辟的论述。孔子说:"学而不思则罔,思而不学则殆。"③孟子不仅区分感性认识和理性认识(思维),而且深刻论述了两者的关系。他说:"耳目之官不思,而蔽于物。物交物,则引之而已矣。心之官则思,思则得之,不思则不得也。此天之所与我者。先立乎其大者,则其小者弗能夺也。此为大人而已矣。"④

概念、判断和推理是思维的基本形式。人类能够产生和使用抽象概念,这是思维的起点。机器能够根据定义去使用某种概念,但它是否能够自己产生和使用抽象概念呢?——这是一个问题。类似地,人类能够用概念来做判断,用判断来进行推理,那么,计算机是否也具有这种能力呢?显然,判断和推理也是人工智能的重要标准。

1997 年 5 月 11 日,IBM 的计算机程序"深蓝"击败了世界排名第一的国际象棋大师加里·卡斯帕罗夫。就在几天前,2016 年 3 月 15 日,围棋"人机大战"第五场,经 5 个小时激战,韩国棋手李世石九段不敌 AlphaGo,输掉最后一场。AlphaGo 以 4:1 总比分取得胜利。此前第 4 场比赛时,李世石找到并利用 AlphaGo 的弱点,曾扳回一局。

此次 AlphaGo 战胜李世石,与 19 年前"深蓝"战胜卡斯帕罗夫有一些重要的区别。主要是围棋的变化更多,更复杂。所以 AlphaGo 采用了学习策略和一些更加高级的算法。而"深蓝"只是按照事先设计的程序办事。这样就让人觉得,击败卡斯帕罗夫的并不是"深蓝",而是设计软件的人;但击败李世石的却是 AlphaGo,是人工智能,是它自身通过学习增长智慧而最终战胜了人。因此可以得出结论,人工智能已经超过了人类智能。事实是否真的如此呢?

毫无疑问,在以计算和推理为主要思维和决策方式的国际象棋和围棋领域,人工智能已经超过人类智能。不仅如此,在以计算、推理和机械行为为特征的其他很多领域,如生产线上的机器人,人工智能也做得比人类更好,更准确,更有效。奇怪的是,在 19 年前"深蓝"战胜卡斯帕罗夫的时候,在数十年前机器人普遍取代人工的生产线日益成为现代工业的生产方式的时候,人们并没有惊呼人工智能战胜了人类智能,为何这一次 AlphaGo 战胜李世石的时候人们却发出这样的惊呼?

事实上,这一次 AlphaGo 战胜李世石和以前"深蓝"战胜卡斯帕罗夫并无本质的区别,与生产线上机器人代替人工也并无本质的区别。在某些需要特殊技能的领域,机器可以和人类做得一样好,甚至超

① http://www.psychology4all.com/Thinking.htm.

② Searle, John R. (2004) *Mind: A Brief Introdmtion.* New york: Oxford University Press, pp.107 – 132.

③ 《论语·为政》。

④ 《孟子·告子上》。

过人类。但并不能由此得出普遍的结论:人工智能已经超过人类智能。在一些看似简单,但需要直觉、灵感、顿悟和创造性思维的领域,如面孔识别、直觉判断、交感交流、创新思维,人工智能的水平甚至不如一个婴儿。至于将已有的知识进行综合创新,如阿基米德在洗澡时灵光一闪发现浮力定律、牛顿看见苹果落地而发现万有引力、爱因斯坦通过计算加猜测而得出著名的质能公式:$E = mc^2$,则是人工智能可望而不可即的。

世界著名心智和语言哲学家塞尔(John R. Searle)在 20 世纪 80 年代提出的"中文房间"人工智能模型,迄今并无任何计算机能够通过这个模型的测试。① 就是说,数字计算机并没有真正意义的人类智能,而且永远也不会有。② 塞尔的这个著名论断并未过时。

(五)文化层级的认知

文化认知是五个层级中最高层级的认知形式,它也是人类特有的认知形式。总结来说,人类认知就是以语言为基础,以思维和文化为特征的高阶认知。文化是人所创造的一切对象的总和,是人的创造物,包括物质存在、社会存在和精神存在。

广义地说,科学、艺术、哲学和宗教都属于文化的范畴。科学、艺术、哲学和宗教从不同的角度,在不同的层次上反映了人类心智,反映了人类对物质世界和精神世界的认识。

在文化认知的领域,似乎仍然是人类智能的一统天下,机器和人工智能无法问津。中医诊断的专家系统已经开发出来了,但我相信多数人包括我自己都不会相信这个专家系统胜过相信专家。写格律诗的计算机软件也开发出来了,并且通过大数据方法能够模仿某位诗人如李白写一首格律诗来祝贺朋友的生日,但我相信这样的李白诗篇恐怕连赝品都算不上。一幅张大千的画或一方齐白石的印可能价值连城,而一幅计算机模仿的张大千的画或一方计算机雕刻的齐白石的印则不值几文钱。曹雪芹是独一无二的,因而《红楼梦》也是独一无二的,没有人会把高鹗的后四十回续书看得与曹雪芹的前八十回有相同的文学价值。心智是涉身的,曹雪芹的心智和他个人的特殊的人生经验决定他才是曹雪芹,也决定他写的《红楼梦》才是《红楼梦》,他的特殊的心智和认知方式是任何人工智能也无法模仿的。

四、几点结论和简单讨论

最后我们给出几点结论并作简单的讨论。

(一)人工智能是在不断进步,不断提高,不断发展的。

自有计算机就有人工智能,人工智能的历史和计算机的历史一样久远。自 20 世纪 50 年代以来,人工智能的理论和技术在不断发展,不断提高,人工智能的技术在不断进步,这是事实。刚刚结束的李世石 PK 谷歌狗 AlphaGo 的人机大战,AlphaGo 以 4:1的成绩战胜李世石,这个战绩比 19 年前"深蓝"战胜卡斯帕罗夫要辉煌得多,一是因为围棋的复杂程度远胜于国际象棋。围棋 19×19 的方阵,共有 361 个落子点,所以整个围棋棋局的总排列组合数高达 10 的 171 次方,据说这比全宇宙的原子数 10 的 75 次方还要多得多! 而计算机的复杂性对人工智能的算法提出了挑战;二是因为围棋的规则非常简单,这就使得过去人工智能常用的靠规则和套路和运算速度取胜的人工智能策略失去优势,必须设计新的能够在围棋博弈中取胜的新策略。事实上,此次的 AlphaGo 使用了很多全新的策略,如深度学习策略、类神经网络系统、价值评估策略等等,这才是 AlphaGo 致胜的关键,也是从"深蓝"到 AlphaGo 人工智能理论发展和技术进步之所在。

(二)在人类认知的所有五个层级上,人工智能都是在模仿人类智能,在总体上并未超过人类智能。

在神经、心理、语言、思维、文化等人类认知各个层级上,人工智能都是在模仿人类智能。这五个层

① 蔡曙山:《哲学家如何理解人工智能》,《自然辩证法研究》2001 年第 11 期。

② 蔡曙山:《关于哲学、心理学和认知科学的 12 个问题与塞尔教授的对话》,《学术界》2007 年第 3 期。

级的人类认知,是人类心智进化各阶段在人的脑与认知系统中保有的能力和智能方式。因此,五个层级的人类心智和认知也就是五个层级的人类智能。本文从人类心智和认知的五个层级考察了人工智能与人类智能的差异,可以看出,在人类心智和认知即人类智能的各个层级上,人工智能都是在模仿人类智能,并且都未能达到人类智能的水平;越是较高层级的认知,人工智能越是逊于人类智能,特别是在高阶认知这个层级上,即在语言、思维和文化层级上,目前人工智能是远逊于人类智能的。事实上,在高阶认知这个层级上,人工智能和人类智能这两种智能方式是截然不同的。

(三)机器学习也只是对人类认知能力的一种模仿,不能作过高的评价。

这次人机大战和机器胜出,最受人们称道的就是 AlphaGo 的深度学习策略,包括通过网络架构与大量样本预测对手落子的神经网络策略、通过价值评估计算胜率的评价网络,以及根据有限选项中计算最佳解的蒙地卡罗搜索树。AlphaGo 就是根据这三个函数来找出最佳动作。

问题是 AlphaGo 这台先进的机器,其学习能力究竟处于什么水平呢? 是否已经出现了超过人类,甚至要终结人类的机器? 首先,AlphaGo 的胜利是数学和逻辑的胜利,说明在某些特定领域的数学计算和逻辑推理方面,机器有可能胜过人类的左脑。但 AlphaGo 并不是真的理解了什么是围棋,更不能将下围棋当作一种艺术来享受,正如计算机可以惟妙惟肖地模仿人类的消化过程,但它却不可能去品尝和享用一块汉堡包或一份中式大餐一样。在心理与直觉方面,机器远远不如人类的右脑。例如,欣赏音乐和绘画,在千万人中,"蓦然回首,那人却在灯火阑珊处"。在这些方面,计算机远不如人类。这种差别不是数量级的差别,而是本质的差异。其次,人和动物都具有学习能力。甚至一些低级动物也具有学习的行为,例如老鼠走迷宫、鹦鹉学舌等等。利用狗的灵敏的嗅觉,经过学习和训练,狗可以完成查找毒品和爆炸物这类特殊的任务,其在特定领域的专业能力可以胜过人类。实际上,这种智能行为只需要刺激反应和记忆就可以了,属于神经和心理层级的认知,是比较低级的认知形式。至于比较高级的认知形式如语言认知、思维认知和文化认知,连一些非人类的高级动物如灵长动物黑猩猩都不具备,更遑论完全不具备心智的冰冷的机器。其三,人工智能就是一种常规的技术进步,机器学习也只是对人类认知能力的一种模仿,即使机器在某些专业领域的能力超过人类,也不能作过高的评价,更不必引起恐慌。在技术进步史上,超过人类能力的机器很多,汽车跑得比人快,火车拉得比人多,飞机飞得比人高……"人类跑不过汽车的时候为何没有那么恐慌呢? 跑步这项运动到现在也好好的,奥运金牌也不是都被法拉利拿走了……,所以真的不必太过紧张。"①至于说机器具有人类情感,会谈恋爱和生育后代,如果不是好莱坞的科幻电影,就只有是一些人的无知梦幻或天方夜谭。

(四)机器永远不会具有人类智能,因此在可预见的未来也不会出现超越或控制人类的机器。

著名的塞尔人工智能理论"中文房间论证"并未过时,根据这个论证,数字计算机(digital computer)永远不会具有智能。②

按照弱人工智能(Artificial Narrow Intelligence)和强人工智能(Artificial General Intelligence)的划分,目前的人工智能通通属于弱人工智能,即让机器具有某种智能的行为(can machine act intelligently)。尽管在某些特殊的领域,专家系统已经达到甚至超过人类智能,例如 19 年前战胜卡斯帕罗夫的"深蓝"和刚刚战胜李世石的 AlphaGo,也包括在当代工业生产线上广泛使用的各种专业机器人。但这些人工智能包括"深蓝"和 AlphaGo 的人工智能都只是属于弱人工智能。而强人工智能,即让机器能够真正地像人类一样地思考(can machine really think),目前只是出现在好莱坞的科幻电影里或强人工智能者的信

---

① 尹相志:《浅谈 AlphaGo 所涉及的深度学习技术》,http://learning.sohu.com/20160316/n440632177.shtml。

② Searle, J., 1980, 'Minds, Brains and Programs', *Behavioral and Brain Sciences*, 3: 417 - 457; 1984, *Minds, Brains and Science*, Cambridge, MA: Harvard University Press. The Chinese Room Argument, *Stanford Encyclopedia of Philosophy*. http://plato.stanford.edu/entries/chinese-room/

仰里。迄今为止,并没有任何一部真正理解人类语言的机器,因此更不会有像人一样能够进行创造性思维的机器和具有人一样的文化生存方式的机器。

那么,人是否可以像上帝那样,创造具有心智的生命呢? 答案是肯定的。事实上,已经有疯子或者"科学狂人"试图扮演这样的角色。美国"科学坏蛋"(bad man of science)文特(J. Craig Venter)已经在他的实验室合成生命。① 但这并不是机器人,而可能是有心智的人工生命(对它的讨论已经超出本文范围,适当时我们可以再讨论)。可惜的是,在 AlphaGo 战胜李世石以后一些人鼓噪的"机器战胜人类"的喧嚣声中,人们对真正的人类命运问题应有的关注被完全转移了。

Artificial Intelligence and Human Intelligence: AView on Human-Computer Competition from the Five-Level Theory of Cognitive Science

Cai Shushan[1] Xue Xiaodi[2]

(1. *College of National Culture and Cognitive Science*, *Guizhou Minzu University*, *Guiyang*, *Guizhou* 550000, *China*; 2. *Department of Psychology*, *Tsinghua University*, *Beijing* 100084, *China*)

Abstract: The present authors hold that human mind and cognition can be viewed at five levels, i. e. , at the levels of neuro, psychology, language, thinking and culture. This paper argues that artificial intelligence (AI, for short) simulates human mind at all the five levels of human cognition. On the whole, AI has not surpassed human mind although it has kept making progress. So far, AI still lags far behind human intelligence at the higher-levels of cognition, that is, at the levels of language, thinking and culture. In fact, AI and human intelligence are two entirely different modes of intelligence. Machine learning is only a simulation of human cognitive ability. It is unnecessary for us to overestimate it or to be scared of it. What is more, we can by no means artificially create panic. Chinese Room Argument, a famous AI model and standard put forward by John R. Searle, is not out of date. According to this argument, digital computer will never have human intelligence. Therefore, no artificial intelligence that may surpass human intelligence will appear in the foreseeable future.

Key words: human mind, human cognition, human intelligence, artificial intelligence (AI), cognitive science

(责任编辑 李 铄)

① Craig Venter, From Wikipedia, the free encyclopedia. https://en.wikipedia.org/wiki/Craig_Venter